CMTS Upstream Utilization Distribution

John J. Downey – Cisco System
CMTS Technical Leader
4/17/17

Customers like to see CMTS upstream (US) utilization relatively equal and balanced amongst USs in a fiber node or US connector. The question to ask is: "Why is it absolutely critical to have perfect balance of US utilization?" Is it because of perception and/or tools? Do customer thresholds cause flags in reporting and alarms?

A case could be made that as long as an US is not congested, then moving CMs around is not practical or efficient. The CMTS MAC scheduler assigns minislot allocation from US0 through the rest of the USs in the MAC Domain (MD) [interface cable]. For non-US bonded CMs, if an US is congested, then US load balancing can be implemented based on utilization.

When US bonded traffic makes US0 congested, then the other USs will get assigned minislots from the MAC Scheduler in a round-robin fashion. This could influence D1.1/2.0 CM movement, which isn't optimum. The bigger issue could be reporting and customer perception.

When the first US in a bonding group approaches congestion, the next numerical US will get assigned more grants. Many times the utilization could look like this:

- ```
  cbr8#sh int c1/0/0 mac-scheduler | in uti
      Avg upstream channel utilization : 87%
      Avg upstream channel utilization : 65%
      Avg upstream channel utilization : 53%
      Avg upstream channel utilization : 30%
  ```

The code by default assigns minislots starting with the first US in the bonding group. Some would argue that the default behavior of the MAC Scheduler is adequate as it keeps the minislots in a single US for better latency/jitter concerns.

Note: Originally, a best practice was to make sure the first US in the BG was **NOT** a smaller channel like TDMA-ATDMA and bonded with multiple ATDMA channels.

In response to customer concerns, Cisco came out with a feature to evenly distribute the D3.0 minislot allocation across all the USs the CM uses in a bonding group. The cable interface command; `cable upstream balance-scheduling` was introduced and uses an evenly distributed allocation. This works very well when all the USs in the mac domain are used in a single US BG. The problem is:

- Some CMs use single-ch bonding (possibly multiple instances)
- Some CMs could be in partial mode
- Some CMs may be doing 2-ch bonding (subset BG)
- There may be architectures consisting of 2 physical fiber nodes in a service group

Note: Even if not perfect, it's still suggested to use the above command.

An enhancement to this command was suggested to just look at the BG used by the CM, but it turned out to be more involved than originally believed. Bonding is done at the service flow level and not really the CM. This means a CM could be assigned to a 4-ch Transmit Channel Set (TCS), but flows like VoIP [UGS] that are scheduled will use one of the default, single-channel BGs. Best Effort (BE) data flows will use the largest available BG it can support by default.

Note: The uBR10k activated dynamic US load balance based on utilization for D3.0 CMs between BGs [12.2 SCI], but it was never implemented on the cBR-8. In most cases it's not really needed as not many systems have enough US spectrum to do 2, 4-ch US BGs.

If the current US utilization is not relatively balanced, there are a few things to look for and consider:

1. Is the cable interface command; `cable upstream balance-scheduling` configured?
   a. It is not on by default.

2. Is the command configured, but not actually working?
   a. A recent issue has arisen and DDTS filed: CSCvd98223 - "`cable upstream balance-scheduling`" is configured, but "`us_balance:OFF`".
   b. Verify with the mac-scheduler command:
      i. `sh int c1/0/0 mac-scheduler | in balance`
      ```
      us_balance:ON
      us_balance:ON
      us_balance:ON
      us_balance:ON
      ```
   c. We are still researching why sometimes it's becoming deactivated even though it's configured. It may be related to SUP or linecard failovers.
   d. The work-around if "off" is to do: "`no cable upstream balance-scheduling`" and then reconfigure; "`cable upstream balance-scheduling`".

3. Does the architecture consist of two physical fiber nodes (FNs) per MAC Domain (MD)?
   a. Due to the US minislot allocation working on a round-robin scheme, this makes the `cable upstream balance-scheduling` not "perfect" when more than 1 US BG is in the MAC Domain (interface cable).
      i. Work-around 1: Configure 1 MAC Domain per physical fiber node.
         1. But this creates its own issues, such as: less D2.0 LB and complexity.
      ii. Work-around 2: Make US BGs with staggered USs. Assuming a 2-FN MD, it could be designed like this:
      ```
      • interface Cable1/0/0
         load-interval 30
         downstream Integrated-Cable 1/0/0 rf-channel 0-15
         upstream 0 Upstream-Cable 1/0/0 us-channel 0
         upstream 2 Upstream-Cable 1/0/0 us-channel 1
         upstream 4 Upstream-Cable 1/0/0 us-channel 2
         upstream 6 Upstream-Cable 1/0/0 us-channel 3
         upstream 1 Upstream-Cable 1/0/1 us-channel 0
         upstream 3 Upstream-Cable 1/0/1 us-channel 1
         upstream 5 Upstream-Cable 1/0/1 us-channel 2
         upstream 7 Upstream-Cable 1/0/1 us-channel 3
         cable upstream bonding-group 1
           upstream 0
           upstream 2
           upstream 4
           upstream 6
         cable upstream bonding-group 2
           upstream 1
           upstream 3
           upstream 5
           upstream 7
      ```

4. A new algorithm will be introduced for the `cable upstream balance-scheduling` command in IOS-XE 16.6.1 for better US traffic distribution. This is tracked via CSCvc74614: When 2 BGs are setup, US LB does not work. This adds the advantage of the following as well:
    a. Potentially no, or less, need for D3.0 dynamic utilization LB between US BGs
    b. Less D2.0 LB movement
    c. Would still work with single-channel BG flows along with embedded/overlapping BGs
    d. Will work with different architectures
    e. Another DDTS CSCuw05662

Tip: Suggest registration-based, modem-count LB, then the configured command `cable upstream balance-scheduling` will use the new upstream algorithm to keep utilization balanced.

US Utilization Verification

When verifying US utilization, there are a few commands typically used. It must be understood how the reporting is done. Is it averaged over samples, and/or time, or instantaneous?

The mac-scheduler command is frequently used to see utilization, but it is an instantaneous reading and not a good indicator of balanced traffic. It is on a .5 sec update and based on minislot allocation, not actual layer 3 traffic. It includes all overhead and a true indicator of what some call "channel utilization" vs "bandwidth utilization" where you look at bps and divide by usable rate for the channel.

- `cbr8#sh int c1/0/0 mac-scheduler | in uti`
  ```
  Avg upstream channel utilization : 36%
  Avg upstream channel utilization : 30%
  Avg upstream channel utilization : 41%
  Avg upstream channel utilization : 37%
  ```

Another way to see utilization would be to take advantage of the LB reporting. The DS utilization is on a rolling, time avg window based on the `load-interval` setting under the integrated, modular, and/or wideband interfaces. The default is 300 sec, but it is suggested to use the lowest setting of 30. The US utilization reporting under the load balance command was based on the load-interval under the cable interface, but may be internally set to 30 sec now. It is still advisable to set this load interval to be safe.

- `interface Cable1/0/0`
    `load-interval 30`

D3.0 LBGs are created automatically when fiber nodes are configured and will be shown as below:

- `cbr8#sh cab load all | beg load`
```
Current load:
DOCSIS load-balancing load
Interface          State    Group       Utilization  Rsvd NBCM  WB/UB Weight
                                                           Total Total
In1/0/0:0(645 MHz)  initial 2147508224 61%(61%/38%)  0%   0     11    37
Us1/0/0:0           initial 2147508224 53%           0%   0     13    30.7
Us1/0/0:1           initial 2147508224 25%           0%   0     13    30.7
Us1/0/0:2           initial 2147508224 43%           0%   0     13    30.7
Us1/0/0:3           initial 2147508224 30%           0%   0     13    30.7
```

Another reporting tool would be via snmp/MIBs and the preferred method for most operators. A global command exists to set the rolling window.

- `cable util-interval x`

A MIB and OID can be used as well: `"1.3.6.1.2.1.10.127.1.3.9.1.3"`,
`//docsIfCmtsChannelUtUtilization`